Genome-wide mapping of global-to-local genetic effects on human facial shape

Supplementary Notes, Figures and Tables

Supplementary Note 1: Selected candidate genes in 15 replicated GWAS regions

A literature-based annotation of genes, located within 500kb up-and downstream from the 15 replicated SNPs was performed and the involvement of these genes in facial variation was investigated. Of particular interest were regions previously associated with syndromes with a known facial phenotype or previously described in facial GWAS studies.

rs72691108:G>A at 1p12 is located 300kb upstream from *TBX15* (T-Box 15, a transcription factor), deficiency of which causes Cousin syndrome[1] (OMIM #260660). Among other symptoms, Cousin syndrome has a specific facial phenotype characterized by frontal bossing, narrow palpebral fissures, hypertelorism and low-set and posteriorly rotated ears. The effect associated with rs72691108:G>A (1p12, Figure 3) was located in the upper facial quadrant, with a depression (blue color) of the forehead laterally and protrusion centrally, recapitulating the frontal bossing phenotype.

The *HOXD* (homeobox D) cluster is a highly-regulated transcription factor family, known to play an important role in morphogenesis in all multicellular organisms[2]. rs970797:T>G in 2q31.1 is located downstream of several *HOXD*-genes (*HOXD1,3,4,* and *HOXD8-13)*. Expression of these genes in the pharyngeal arches provides evidence for their involvement in craniofacial development[3]. Interestingly, this locus was one of the two results that affected more than one facial quadrant of the dendrogram (Figure 3), possibly reflecting the important and diverse role of this cluster in facial morphogenesis.

The role of *PAX3* (Paired Box 3, a transcription factor) in facial morphology, and more specifically in nasal shape, has been identified and described in several independent facial GWAS studies[4-6]. *PAX3* encodes a transcription factor, important in fetal and neural crest development as a regulator of cell proliferation, migration and determination[7]. Mutations in *PAX3* cause Waardenburg syndrome type I (OMIM #193500). One of the major phenotypic characteristics of this syndrome is dystopia canthorum, the lateral displacement of the ocular inner canthi. From Figure 3, it was observed that the effects of rs10176525:A>T in 2q36.1 are unmistakably related and focused on the nose quadrant.

rs9995821:C>T in 4q31.3 is located 250kb upstream of *DCHS2* (dachsous cadherin-related 2), which was described as a candidate gene for columella inclination[6]. *DCHS2* is involved in a regulatory network controlling cartilage differentiation and polarity during vertebrate craniofacial development[8]. This network also includes *SOX9* (SRY box 9, coding for a DNA binding protein). Mutations in the latter gene lead to Campomelic Dysplasia (OMIM #114290), associated with the long-bone defects and a set of craniofacial malformations termed Pierre Robin Sequence (PRS) characterized by cleft palate, micrognathia, and flat face[9]. Interestingly, mutations

within a large gene desert located upstream from the *SOX9* gene in some cases give rise to the isolated PRS and have been proposed to affect long-range craniofacial enhancer function[10]. We identified and replicated two variants falling within this upstream gene desert, rs5821892:C>CG (17q24.3 (b)) in the vicinity of *SOX9* as well as rs11655006:A>G (17q24.3 (a)) in the vicinity of a non-coding transcript, *BC039327/CASC17*, a breakpoint upstream of *SOX9*[11]. Both SNPs showed effects on nasal shape, but with a different global-to-local focus.

The genes *DLX5* and *DLX6* (distal-less homeobox 5 and 6, transcription factors) are located 500kb downstream of rs10238953:G>A (7q21.3). In humans, *DLX5/DLX6* are osteoblast-specific genes and disruption of these genes is associated with split hand/foot malformation with/without craniofacial dysmorphology (OMIM #183600), but also with other craniofacial anomalies such as cleft palate[12]. Furthermore, *DYNC1I1* and *SHFM1* are also located in the same chromosomal region, both known to be involved in split hand/foot malformation[13], indicating that this locus is crucial for normal development. The pattern of *DLX* gene expression in mice gives rise to the proximodistal axes of the jaws[14] and is involved in muscle formation in this region[15], indicating a key role for *DLX6* in the development of jaw structure. This is supported by the findings in our study, where rs10238953:G>A is associated with morphological changes in the mandibular region.

Supplementary Note 2: Integration with facial GWAS literature

To date, six facial GWAS starting from 2D and/or 3D facial images have been reported[4-6,16-18] including a total of 33 SNPs, all of which achieved at least nominal genome-wide significance in the respective studies either at discovery or meta-analysis stage. First, in an effort to provide additional validation of our approach, we tested a subset of these SNPs in our discovery PITT cohort, using the 63 facial segments as a proxy for the measures (typically linear distances) used in these various studies. Second, in an effort to provide additional support (besides the PSU replication) for our discovered loci listed in Table 1, we consulted GWAS central (www.gwascentral.org) and Phenoscanner[19] for facial results, which holds less than genome-wide significant results from previous studies. Third, we cross-referenced our loci (and not just the peak SNPs) with a recent study[20] based on self-reported information on 42 human traits using questionnaires, including two questions on facial morphology, one about the presence/absence of a chin dimple and one about nose size. Loci involved in any of the three integration efforts are indicated in Table 1.

For our first integration effort, a total of 16 out of the 33 previously reported SNPs, were selected for testing as listed in Table S3. Four SNPs from Cole et al.[16], were omitted, because three of them were associated with facial size, a trait that was specifically corrected for in this work, and one of them was below our MAF cut-off of 1%. 13 SNPs from Shaffer et al.[17] and their follow-up study[18] were omitted, because the same discovery cohort was used as the present study, and replication could therefore not be claimed. Our FDRd and Bonferroni significance thresholds for replicating SNPs from the literature were determined as 1.59×10^{-4} and 8.01×10^{-5}, respectively. Polar dendrograms and heatmaps of facial displacement for the SNPs in Table S3, are given online (https://www.esat.kuleuven.be/psi/research/global-to-local-

facial-phenotyping). Authors P.C. and J.R. visually inspected the genetic effects in at least nominally ($p \leq 0.05$) associated facial segments, judging them independently as similar to the previously reported associated facial traits or not. Subsequently, our replication results were categorized and graded as A, B, C or F based on effect similarity and p-value. 10 of the 16 SNPs were categorized as A or B, representing good replication and only three SNPs were categorized as F, representing complete failure to replicate.

Interestingly, five of the 16 previously published SNPs represented four loci observed in our GWAS (*PAX3, PAX1, DCHS2, SUPT3H/RUNX2*). Furthermore, our results for *PAX3*, the strongest facial GWAS gene to date which has been identified in three independent studies[4-6], were equally strong and convincing. One previously reported SNP (rs7559271:G>A, in Table S3) in *PAX3* was also well below our study-wide significance ($\leq 1.28 \times 10^{-9}$). Within proximity, another *PAX3* SNP (rs10176525:A>T, Table 1), showed the strongest statistical evidence, to date, and was also replicated in the PSU cohort. Also of note was rs2045323:A>G which was associated with three closely related nose features (columella inclination, nose protrusion, and nose tip angle) by Adhikari and colleagues[6]. Our results for this particular SNP are depicted in Figure S4, were we observed the same facial effects as rs9995821:C>T listed in Table 1 and shown in Figure 3. These two SNPs are located in close proximity (< 250 kb) on 4q31. More importantly, the facial segment most strongly associated in the GWAS was the nose segment focusing on the columella and its integration with the nose tip, which is very similar to the traits reported in the primary publication.

For our second integration effort, we looked up results for the SNPs identified in our discovery efforts (Tables 1 and S1) in the facial GWAS by Adhikari et al. [6] and Paternoster et al.[4], on GWAS Central and Phenoscanner. Nine of our discovery SNPs were available, and all showed evidence of replication (Table S4). The traits reported, were closely related to the genetic facial effects we observed (Figure 3, where blue/red indicates a decrease/increase in local facial prominence). For example, 17q24.3 and 2q36.1 show a depression and protrusion (respectively) of the nose, 6p21.1 and 3q21.3 depict a widening of the nose bridge/alae, 4q31.3 and 3q27.1 both change the inclination of the columella (but 4q31.3 more directly than 3q27.1 where the strongest effect was on the nose tip), and finally 2p21 affects the shape of the chin. Interestingly, locus 20p11.22 (rs2424390:T>A), which replicated nominally ($p=2.47 \times 10^{-2}$) in our PSU cohort, and was therefore not depicted in Figure 3, showed a strong replication in Adhikari et al. Figure S5 shows that the effect of this SNP in the most strongly associated facial segment is related to nasal alae breadth. Only the similarity between 2q31.1 and PC 11 in Paternoster et al. is not straightforward. Note that these facial effects observed, support our observational exercise to define similarity to previously associated traits using A to F categorizations in Table S3.

For our third integration effort, we found significant effects for 10 out of the 18 loci, using a 500kb window around each peak SNP, from Table 1 in Pickrell et al.[20] as reported in Table S5. These loci are primarily (with the exception of 1p12 and 2q31.1) affecting nose and/or chin segments as observed in Figure 2 and 3. Therefore, our findings were generally consistent with Pickrell et al., but with the additional novelty that we provided much more detailed descriptions of their facial effects. For example,

1q31.3, 2p21, and 7q21.3 affect the chin in different ways at different scales. It is also not easy to compare the presence of a chin dimple (a nonmetric trait) with the effects shown in this work. 1p12 and 2q31.1 are potentially loci that pleiotropically affect multiple independent traits, which is interesting both mechanistically and evolutionarily.

The results on seven SNPs that were previously identified in Shaffer et al.[17] using the same discovery PITT cohort are summarized in Table S6 and depicted online (https://www.esat.kuleuven.be/psi/research/global-to-local-facial-phenotyping). Two of the three SNPs that were discovered in Shaffer et al. showed strong and similar trait related associations in this study, but failed to reach the genome-wide discovery threshold. On the other hand, for one SNP (rs2424399:A>C) of the four that were only identified after a meta-analysis with an additional cohort in Shaffer et al., we achieved genome-wide significance using the PITT cohort only. Together with rs2424390:T>A (Table 1), this SNP was located in the region (20p11.22) that replicated using the results on GWAS central. This locus was annotated with *PAX1* (Paired Box 1, a transcription factor), known to be involved in chondrocyte differentiation[21]. Mutations in *PAX1* can cause otofaciocervical syndrome, which is characterized by a sunken nasal root and excessive nasal narrowing[22], which are consistent with our results for this region. Finally, a combined listing of all facial GWAS outcomes from images is provided in Table S7.

ANNOTATION					EFFECT		DISCOVERY							REPLICATION				
Region	SNP	Gene	a	A	Quadrant	Module	CC	CC2	DF1	DF2	p-value	MAF	Nr SNPs	B	SE	p-value	MAF	nr SNPs
1p32.1	rs4916068	intergenic	C	T	2	41	0.16	0.03	9	2315	4.81E-10	0.500	28	0.25	0.03	1.39E-12	0.494	28
1p21.3	rs185192810	*CNN3*	T	C	3	54	0.16	0.03	12	2285	2.14E-08	0.012	2	0.12	0.11	2.77E-01	0.012	0
1p12	rs142713255	*TBX15*	A	G	3	54	0.16	0.03	12	2187	3.09E-08	0.012	1	0.20	0.12	8.03E-02	0.012	0
1p12	rs72691108	*TBX15*	A	G	4	7	0.25	0.06	41	2242	5.81E-14	0.236	224	0.34	0.04	1.01E-15	0.243	224
1q31.3	rs2821116	*ASPM*	A	T	3	13	0.24	0.06	21	2307	2.63E-19	0.210	342	0.32	0.04	2.29E-15	0.209	342
2p21	rs6740960	*PKDCC*	T	A	3	26	0.21	0.05	14	2053	3.03E-14	0.497	2	0.17	0.04	3.44E-06	0.499	2
2p16.2	rs11690122	*ASB3, CHAC2, ERLEC1, GPR75*	A	C	4	28	0.19	0.04	23	2257	1.04E-08	0.020	1	0.08	0.12	5.42E-01	0.018	0
2q12.1	rs17632157	*POU3F3, TGFBRAP1, FHL2*	T	C	2	5	0.20	0.04	25	2214	4.21E-10	0.020	1	0.03	0.10	7.37E-01	0.024	0
2q31.1	rs970797	*HOXD* cluster	T	G	2,3	38	0.17	0.03	8	2320	6.17E-11	0.434	15	0.11	0.03	1.12E-03	0.428	4
2q36.1	rs10176525	*PAX3*	A	T	2	10	0.21	0.05	20	2308	3.76E-14	0.224	38	0.26	0.04	4.39E-11	0.233	38
3p22.3	rs4955237	*DYNC1LI1*	T	C	3	55	0.16	0.03	12	2316	3.02E-08	0.217	1	0.03	0.04	3.99E-01	0.216	0
3p14.1	rs1523446	*ADAMTS9*	C	T	4	58	0.16	0.03	12	2278	4.27E-08	0.265	1	0.01	0.04	7.34E-01	0.268	0
3q13.31	rs149210078	*LSAMP*	G	A	4	60	0.17	0.03	15	2253	2.69E-08	0.024	1	0.07	0.11	5.29E-01	0.026	0
3q21.3	rs2977562	*RAB7A, ACAD9*	G	A	2,3	16	0.22	0.05	13	2298	1.39E-17	0.244	389	0.20	0.04	8.63E-07	0.248	374
3q27.1	rs58022575	*EPHB3, DVL3*	G	GA	2	20	0.17	0.03	13	2299	7.99E-10	0.473	16	0.21	0.03	2.39E-09	0.467	16
4q31.3	rs9995821	*DCHS2*	C	T	2	45	0.22	0.05	13	2315	3.67E-18	0.200	10	0.39	0.04	7.33E-19	0.204	10
5p14.3	rs62354288	*CDH18*	T	A	2	21	0.17	0.03	12	2032	8.05E-09	0.355	4	0.02	0.04	7.04E-01	0.357	0
5q35.3	rs2913791	*NHP2, ZNF354A*	A	G	3	55	0.16	0.03	12	2304	4.96E-09	0.429	5	0.00	0.03	9.14E-01	0.417	0
6p21.1	rs227833	*SUPT3H*	C	G	2	10	0.21	0.05	20	2307	3.38E-14	0.261	164	0.14	0.04	5.63E-04	0.254	150
6q23.2	rs5880172	*RPS12, EYA4*	A	ACT	4	30	0.21	0.04	24	2267	6.10E-12	0.312	33	0.26	0.04	6.91E-13	0.297	33
6q27	rs56060091	*MILT4, SMOC2*	G	C	2	11	0.18	0.03	17	2145	1.90E-08	0.490	1	0.01	0.04	7.29E-01	0.486	0
7q21.3	rs10238953	*DLX6, DYNC1L1*	G	A	3	55	0.24	0.06	12	2316	1.06E-22	0.140	256	0.55	0.05	1.29E-26	0.137	256
8p23.2	rs118029464	*intergenic*	G	A	3	27	0.17	0.03	16	2287	6.90E-09	0.012	2	0.04	0.11	7.49E-01	0.012	0
9q21.2	rs72745682	*PSAT1*	A	G	2	41	0.15	0.02	9	2306	3.28E-08	0.100	1	0.00	0.05	9.94E-01	0.091	0
10q25.1	rs117365954	intergenic	C	G	4	59	0.18	0.03	19	2231	2.31E-08	0.011	1	0.14	0.20	4.74E-01	0.007	0
12q13.13	rs73099938	*KRT18, RARG*	T	C	1	37	0.16	0.03	11	2175	4.82E-08	0.012	1	0.12	0.13	3.66E-01	0.013	0
12q21.31	rs11439959	*ALX1*	CA	C	4	58	0.17	0.03	12	2156	1.89E-08	0.489	1	0.07	0.04	4.89E-02	0.496	0
14q32.11	rs74803822	*CALM1, RPS6KA5, C14ORF159*	C	T	3	27	0.19	0.03	16	2299	1.12E-10	0.011	6	0.04	0.12	7.19E-01	0.011	0
16p11.1	rs62060567	intergenic	A	T	3	26	0.17	0.03	14	2246	3.12E-08	0.106	1	0.03	0.06	6.00E-01	0.113	0
17q24.3	rs11655006	*BC039327*	G	A	2	2	0.27	0.07	33	2288	5.24E-21	0.376	311	0.15	0.03	1.77E-05	0.364	308
17q24.3	rs5821892	*SOX9*	C	CG	2	10	0.19	0.03	20	2293	2.30E-09	0.427	31	0.13	0.04	3.01E-04	0.453	26
18p11.32	rs12326464	*SMCHD1*	A	G	3	54	0.16	0.03	12	2285	7.78E-09	0.012	13	0.25	0.16	1.13E-01	0.006	0
18p11.21	rs199828028	*ZNF519*	T	TG	4	14	0.30	0.09	32	1019	2.14E-08	0.087	2	0.01	0.06	8.67E-01	0.147	0
19q13.11	rs287104	*KCTD15*	G	A	2	41	0.17	0.03	9	2319	1.26E-10	0.349	8	0.16	0.04	2.86E-05	0.329	8
20p11.22	rs2424390	*PAX1*	A	T	2	11	0.17	0.03	17	2291	1.50E-08	0.233	15	0.09	0.04	2.47E-02	0.233	0
Xq21.33	rs146951439	intergenic	C	A	3	54	0.17	0.03	12	2174	1.68E-08	0.013	1	0.07	0.10	5.19E-01	0.015	0
Xq23	chrX:114930965	*PLS3*	T	TGAG	3	27	0.17	0.03	16	2181	2.99E-08	0.018	2					0
Xq28	rs62620964	*FAM58A, SLC6A8, NAA10, HCFC1*	A	T	2	44	0.16	0.03	10	2026	2.79E-08	0.018	1	0.12	0.09	1.86E-01	0.020	0

Supplementary Table S1

Properties of top SNPs in chromosomal regions showing genome-wide significance.

SNP, single nucleotide polymorphism, CC, canonical correlation, CC2, canonical correlation squared, DF1 degrees of freedom numerator, DF2 degrees of freedom denominator, B regression coefficient, SE standard error regression coefficient, MAF, minor allele frequency. Nr SNPs, the number of SNPs reaching genome-wide significance within the same locus. A, major allele, a minor allele. Indicated in red if MAF < 2% or Nr SNPs = 1 or MAF discrepancy between both cohorts. Nr SNPs (Replication phase), the number of SNPs reaching FDRd replication significance within the same locus.

SNP	Region	PITT info	PITT certainty	PITT type	PITT info type0	PSU info	PSU certainty	PSU type	PSU info type0
rs4916068	1p32.1	0.998	0.999	0	-1	0.984	0.988	0	-1
rs185192810	1p21.3	0.89	0.996	0	-1	0.834	0.996	0	-1
rs142713255	1p12	0.736	0.981	0	-1	0.797	0.987	0	-1
rs72691108	1p12	0.988	0.995	0	-1	0.98	0.991	0	-1
rs2821116	1q31.3	1	1	0	-1	0.997	0.999	0	-1
rs6740960	2p21	0.947	0.964	0	-1	0.943	0.961	0	-1
rs11690122	2p16.2	0.914	0.995	0	-1	0.914	0.996	0	-1
rs17632157	2q12.1	0.842	0.989	0	-1	1	1	2	0.848
rs970797	2q31.1	1	1	2	0.989	1	1	2	0.985
rs10176525	2q36.1	1	1	2	0.998	0.99	0.995	0	-1
rs4955237	3p22.3	1	1	2	0.969	0.966	0.983	0	-1
rs1523446	3p14.1	0.989	0.995	0	-1	0.97	0.982	0	-1
rs149210078	3q13.31	0.88	0.99	0	-1	0.804	0.981	0	-1
rs2977562	3q21.3	0.994	0.997	0	-1	0.985	0.992	0	-1
rs58022575	3q27.1	0.997	0.998	0	-1	0.963	0.977	0	-1
rs9995821	4q31.3	1	1	2	0.998	1	1	2	0.984
rs62354288	5p14.3	0.932	0.96	0	-1	0.922	0.955	0	-1
rs2913791	5q35.3	0.998	0.999	0	-1	0.992	0.995	0	-1
rs227833	6p21.1	1	1	0	-1	0.996	0.998	0	-1
rs5880172	6q23.2	0.99	0.994	0	-1	0.982	0.99	0	-1
rs56060091	6q27	0.962	0.974	0	-1	0.932	0.959	0	-1
rs10238953	7q21.3	1	1	2	0.995	0.965	0.989	0	-1
rs118029464	8p23.2	0.901	0.996	0	-1	0.934	0.998	0	-1
rs72745682	9q21.2	0.992	0.998	0	-1	0.976	0.995	0	-1
rs117365954	10q25.1	0.746	0.989	0	-1	0.741	0.991	0	-1
rs73099938	12q13.13	0.686	0.977	0	-1	0.663	0.98	0	-1
rs11439959	12q21.31	0.966	0.978	0	-1	0.972	0.984	0	-1
rs74803822	14q32.11	0.948	0.998	0	-1	0.869	0.996	0	-1
rs62060567	16p11.1	0.949	0.988	0	-1	0.917	0.984	0	-1
rs11655006	17q24.3	0.997	0.998	0	-1	0.998	0.998	0	-1
rs5821892	17q24.3	0.996	0.998	0	-1	0.971	0.98	0	-1
rs12326464	18p11.32	0.905	0.996	0	-1	0.896	0.989	0	-1
rs199828028	18p11.21	0.554	0.832	0	-1	0.848	0.942	0	-1
rs287104	19q13.11	1	1	2	0.986	1	1	2	0.977
rs2424390	20p11.22	0.994	0.997	0	-1	1	1	0	-1
rs146951439	Xq21.33	0.694	0.982	0	-1	0.797	0.989	0	-1
chrX:114930965	Xq23	0.804	0.984	0	-1				
rs62620964	Xq28	0.597	0.955	0	-1	0.628	0.972	0	-1

Supplementary Table S2

Imputation quality measures for the 38 top SNPs listed in table S1.

PITT, refers to the imputations in the discovery cohort, PSU refers to the imputations in the replication cohort. Info, measure of the observed statistical information associated with the allele frequency estimate. Certainty, Average certainty of best-guess genotypes, calculated as the average of the maximum probability across all samples for a given SNP. Type, Internal type assigned to each SNP where type 0 denotes imputed SNPs, type 2 denotes imputation basis (observed) SNPs, and type 3 SNPs are study-only SNPs. Info Type, "Info" quality metric for a type 2 SNP treated as type 0 (i.e. when it was masked); -1 values indicate missing, for imputed (type 0) SNPs

Category	Chromosomal Pos.	SNP	Gene	Orig. GWAS	Orig. Repl.	Repl.	Effect similarity	Reference
A	2q35	rs7559271	*PAX3*	D	Y	8.16E-10	Y	Paternoster et al.
				D	NA	8.16E-10	Y	Adhikari et al.
	4q31	rs2045323	*DCHS2*	D	NA	8.16E-10	Y	Adhikari et al.
	6p21	rs1852985	*SUPT3H/RUNX2*	D	NA	4.59E-07	Y	Adhikari et al.
	1p36.23-p33	rs4648379	*PRDM16*	D	N	1.82E-05	Y	Liu et al.
B	2q35	rs974448	*PAX3*	D	N	2.58E-04	Y	Liu et al.
	5q35.1	rs6555969	*C5orf50*	D	Y	1.28E-03	Y	Liu et al.
	8p14	rs7836044	intergenic	M	N	1.70E-03	Y	Cole et al.
	3p21.1	rs1982862	*CACNA2D3*	D	N	8.52E-03	Y	Paternoster et al.
	4q31	rs12644248	*DCHS2*	D	Y	9.72E-03	Y	Adhikari et al.
	20p11	rs927833	*PAX1*	D	Y	9.84E-03	Y	Adhikari et al.
C	7p13	rs17640804	*GLI3*	D	Y	1.29E-02	Y	Adhikari et al.
	12q21.31	rs10862567	*TMTC2*	D	N	1.42E-02	Y	Paternoster et al.
	2q12	rs3827760	*EDAR*	D	NA	3.37E-02	Y	Adhikari et al.
F	3q28	rs17447439	*TP63*	D	Y	4.65E-03	N	Liu et al.
	5q12.1	rs11738462	*C5orf64*	D	N	3.10E-02	N	Paternoster et al.
	10q24.3	rs805722	*COL17A1*	D	Y	5.60E-02	Y	Liu et al.

Supplementary Table S3

Replication results on previously reported SNPs.

Testing of previously defined genome-wide significant SNPs. A ($p < FDRd=1.59 \times 10^{-4}$ and similar effect), B ($FDRd \leq p \leq 0.01$ & similar effect), C ($0.01 \leq p \leq 0.05$ & similar effect) or F ($p \geq 0.05$ || dissimilar effect). SNP single nucleotide polymorphism. Orig. GWAS, SNP was identified within a single cohort (D) or using a meta-analysis (M), Orig. Repl., SNP replicated (Y), was not tested for replication (NA) or did not replicate (N). Note that SNPs found under meta-analysis required an additional independent test to claim replication. Effect similarity, we observed a similar effect to the reported associated trait (Y) or not (N). References: Paternoster et al. [4], Liu et al.[5], Adhikari et al.[6], and Cole et al[16].

Chromosomal Position	SNP	Candidate Gene	p	Associated Trait	Source
2p21 (42181679)	rs6740960	*PKDCC*	3.03E-04	Chin shape	Adhikari et al.
2q31.1 (177111819)	rs970797	*HOXD* cluster	7.60E-08	pc 11	Paternoster et al.
2q36.1 (223039052)	rs10176525	*PAX3*	1.37E-09	distances n men	Paternoster et al.
3q21.3 (128106267)	rs2977562	*RAB7A, ACAD9*	6.17E-04	Nose wing breadth	Adhikari et al.
3q27.1 (184333169)	rs58022575	*EPHB3, DVL3*	6.78E-04	columella inclination	Adhikari et al.
4q31.3 (154828366)	rs9995821	*DCHS2*	2.14E-06	columella inclination	Adhikari et al.
			4.70E-04	nose tip	Adhikari et al.
6p21.1 (44681840)	rs227833	*SUPT3H*	1.58E-06	nose bridge breadth	Adhikari et al.
17q24.3 (69139583)	rs11655006	*SOX9*	4.50E-05	Nose protrusion	Adhikari et al.
20p11.22 (21626627)	rs2424390	*PAX1*	3.67E-05	Nose wing breadth	Adhikari et al.

Supplementary Table S4

Replication results from web-based GWAS repositories.

Replication of top SNPs in chromosomal regions showing genome-wide significance that were found to affect facial shape in the GWAS central or Phenoscanner databases. SNP single nucleotide polymorphism. p, p-value as reported in the web-based repositories. References: Paternoster et al.[4], Adhikari et al.[6].

Annotation				Chin Dimple			Nose Size		
Chromosome Position	**SNP**	**Gene**	**Facial Segment(s)**	**Position**	**RS**	**p**	**Position**	**RS**	**p**
1p12 (119762175)	rs72691108	*TBX15*	Upper facial quadrant	119452500	rs1766786	1.60E-13			
1q31.3 (197329041)	rs2821116	*ASPM*	Mandible/Chin	197351039	rs2476023	9.70E-39			
2p21 (42181679)	rs6740960	*PKDCC*	Mandible/Chin	42181679	rs6740960	4.30E-13			
2q31.1 (177111819)	rs970797	*HOXD* cluster	Lips/philtrum/nose	177371846	rs59156997	1.70E-19			
2q36.1 (223039052)	rs10176525	*PAX3*	Nose	223529112	rs4674676	1.60E-17	223026935	rs34460569	2.50E-07
3q27.1 (184333169)	rs58022575	*EPHB3, DVL3*	Nose				184339757	rs13097965	1.70E-15
6p21.1 (44681840)	rs227833	*SUPT3H*	Nose	44751668	rs17336368	2.80E-08	44820741	NA	1.20E-08
7q21.3 (96124975)	rs10238953	*DLX6, DYNC1L1*	Chin	96143804	rs11768577	5.30E-53			
17q24.3 (69139583)	rs11655006	*BC039327*	Nose	69123106	rs56347314	5.40E-12			
17q24.3 (70036479)	rs5821892	*SOX9*	Nose				70025587	rs34091987	3.10E-08

Supplementary Table S5

Overlapping loci with self-reported chin dimple and nose size.

Loci from Table 1, that overlap with findings in Pickrell et al.[20] who used self-reported information on 42 phenotypes using questionnaires on very large samples (n > 70.000). Pickrell et al. included two categorical questions on facial morphology, one about the presence/absence of a chin dimple and one about nose size. SNP, single nucleotide polymorphism, Facial Segments, the segments mostly affected by these loci as depicted in Figure 2 and 3. p, p-value as reported in the supplementary material of Pickrell et al.

Region	SNP	Gene	Orig. GWAS	Orig. Repl.	Repl.	Effect similarity	Associated Trait
20p11.22	rs2424399	*PAX1*	M	NA	3.14E-08	Y	Nasal Width
14q11.2	rs8007643	*ZNF219*	D	N	6.45E-05	Y	Nasal Ala
20q12	rs6129564	*MAFB*	D	N	3.03E-04	Y	Cranial base width
Xq13.2	rs11093404	*HDAC8*	D	N	1.69E-03	N	Intercanthal width
1p13.3	rs619686	*ALX3*	M	NA	1.05E-02	N	Intercanthal width
11q22.1	rs12786942	*TRPC6*	M	NA	1.63E-02	Y	Upper facial depth
14q21.1	rs17106852	*PAX9*	M	NA	3.17E-02	Y	Cranial base width

Supplementary Table S6

Testing previously obtained results on the PITT cohort.

Testing of previously defined genome-wide significant SNPs in Shaffer et al.[17] SNP single nucleotide polymorphism. Orig. GWAS, SNP was identified within a single cohort (D) or using a meta-analysis (M), Orig. Repl., SNP replicated (Y), was not tested for replication (NA, which is the case for M) or did not replicate (N). Effect similarity, we observed a similar effect to the reported associated trait (Y) or not (N).

Gene	SNP	Associated Trait	Reference
TMTC2	**rs10862567**	enR.yz: right endocanthion in yz direction)	Paternoster et al, 2012
PAX3	**rs7559271**	nsn-men: nasion to mid-endocanthion point	Paternoster et al, 2012
		nasion position	Adhikari et al, 2016
	rs974448	eyeR-nsn: right eye to nasion	Liu et al, 2012
		eyeL-nsn: left eye to nasion	Liu et al, 2012
	rs10176525	nose quadrant	Claes et al, 2017
CACNA2D3	**rs1982862**	prn-alL: pronasale to left alare	Paternoster et al, 2012
C5orf64	rs11738462	prn-alL: pronasale to left alare	Paternoster et al, 2012
PRDM16	**rs4648379**	prn-alL: pronasale to left alare	Liu et al, 2012
		prn-alR: pronasale to right alare	Liu et al, 2012
TP63	**rs17447439**	eyeL-eyeR: left eye to right eye	Liu et al, 2012
C5orf50	**rs6555969**	zygL-nsn: left zygion to nasion	Liu et al, 2012
		zygR-nsn: right zygion to nasion	Liu et al, 2012
		eyeR-nsn: right eye to nasion	Liu et al, 2012
		eyeL-nsn: left eye to nasion	Liu et al, 2012
COL17A1	**rs805722**	eyeL-nsn: left eye to nasion	Liu et al, 2012
		eyeR-nsn: right eye to nasion	Liu et al, 2012
DCHS2	**rs12644248**	columella inclination (ordinal)	Adhikari et al, 2016
	rs2045323	columella inclination (quantitative)	Adhikari et al, 2016
		nose protrusion	Adhikari et al, 2016
		nose tip angle	Adhikari et al, 2016
	rs9995821	columella/nose tip	Claes et al, 2017
SUPT3H/RUNX2	**rs1852985**	nose bridge breadth (ordinal)	Adhikari et al, 2016
		nose bridge breadth (quantitative)	Adhikari et al, 2016
	rs227833	nose bridge breadth	Claes et al, 2017
GLI3	**rs17640804**	nose wing breadth (ordinal)	Adhikari et al, 2016
		nose wing breadth (quantitative)	Adhikari et al, 2016
PAX1	**rs927833**	nose wing breadth (ordinal)	Adhikari et al, 2016
	rs2424399	nasal width	Shaffer et al, 2016
	rs2424390	nasal width (ala)	Claes et al, 2017
EDAR	**rs3827760**	chin protrusion	Adhikari et al, 2016
SCHIP1	**rs79909949**	centroid size	Cole et al, 2016
PDE8A	rs12909111	allometry	Cole et al, 2016
	rs12908400	allometry	Cole et al, 2016
intergenic	rs7836044	sto-sl: lower lip height	Cole et al, 2016
	rs117438382	Philtrum height	Lee et al, 2017
	rs62031988	nasal width	Lee et al, 2017
	rs113036800	nasal ala	Lee et al, 2017
	rs138440928	unspecified	Lee et al, 2017
	rs4916068	nose tip	Claes et al, 2017
HNRNPR	rs139879053	PC1: upper facial height, midfacial width	Cole et al, 2016
ZNF219	rs8007643	nasal ala length	Shaffer et al, 2016

MAFB	rs6129564	cranial base width	Shaffer et al, 2016
HDAC8	rs11093404	Orbital spacing	Lee et al, 2017
		intercanthal width	Shaffer et al, 2016
ALX3	rs619686	intercanthal width	Shaffer et al, 2016
TRPC6	rs12786942	upper facial depth	Shaffer et al, 2016
PAX9	rs17106852	cranial base width	Shaffer et al, 2016
PARK2	rs9456748	Upper facial height	Lee et al, 2017
FREM1	rs72713618	Central upper lip height	Lee et al, 2017
TBX15	**rs72691108**	Upper face quadrant	Claes et al, 2017
ASPM	**rs2821116**	chin	Claes et al, 2017
PKDCC	**rs6740960**	mandibule	Claes et al, 2017
HOXD cluster	**rs970797**	mouth and philtrum, nose width	Claes et al, 2017
RAB7A, ACAD9	**rs2977562**	philtrum	Claes et al, 2017
EPHB3, DVL3	**rs58022575**	nose bridge	Claes et al, 2017
RPS12, EYA4	**rs5880172**	forehead	Claes et al, 2017
DLX6, DYNC1L1	**rs10238953**	mandibule and chin	Claes et al, 2017
BC039327	**rs11655006**	nose quadrant	Claes et al, 2017
SOX9	**rs5821892**	nose tip, bridge	Claes et al, 2017
KCTD15	**rs287104**	nose tip	Claes et al, 2017
CDH18	rs62354288	nose bridge	Claes et al, 2017
NHP2, ZNF354A	rs2913791	chin	Claes et al, 2017

Supplementary Table S7

Status quo of facial GWAS from 2D or 3D images.

Candidate genes with a lead SNP that reached at least nominal genome-wide significance ($p \leq 5 \times 10\text{-}8$) during discovery or after meta-analysis are listed. SNPs in bold also achieved at least nominal replication ($p \leq 0.05$) in one or more studies. Note that SNPs found under meta-analysis required an additional independent test to claim replication. References: Paternoster et al.[4], Liu et al.[5] Adhikari et al.[6], Cole et al.[16], Shaffer et al.[17], Lee et al.[18]

References

1. Dikoglu, E. *et al.* Homozygosity for a novel truncating mutation confirms TBX15 deficiency as the cause of Cousin syndrome. *American Journal of Medical Genetics Part A* **161**, 3161-3165 (2013).
2. Favier, B. & Dolle, P. Developmental functions of mammalian Hox genes. *Molecular human reproduction* **3**, 115-131 (1997).
3. Minoux, M., Antonarakis, G.S., Kmita, M., Duboule, D. & Rijli, F.M. Rostral and caudal pharyngeal arches share a common neural crest ground pattern. *Development* **136**, 637-645 (2009).
4. Paternoster, L. *et al.* Genome-wide Association Study of Three-Dimensional Facial Morphology Identifies a Variant in *PAX3* Associated with Nasion Position. *The American Journal of Human Genetics* **90**, 478-485 (2012).
5. Liu, F. *et al.* A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS genetics* **8**, e1002932 (2012).
6. Adhikari, K. *et al.* A genome-wide association scan implicates *DCHS2, RUNX2, GLI3, PAX1* and *EDAR* in human facial variation. *Nature communications* **7**(2016).
7. Monsoro-Burq, A.H. PAX transcription factors in neural crest development. in *Seminars in cell & developmental biology* Vol. 44 87-96 (Elsevier, 2015).
8. Le Pabic, P., Ng, C. & Schilling, T.F. Fat-Dachsous signaling coordinates cartilage differentiation and polarity during craniofacial development. *PLoS Genet* **10**, e1004726 (2014).
9. Lefebvre, V. & Dvir-Ginzberg, M. *SOX9* and the many facets of its regulation in the chondrocyte lineage. *Connective tissue research* **58**, 2-14 (2017).
10. Gordon, C.T. *et al.* Long-range regulation at the *SOX9* locus in development and disease. *Journal of medical genetics* **46**, 649-656 (2009).
11. Hill-Harfe, K.L. *et al.* Fine mapping of chromosome 17 translocation breakpoints⩾ 900 kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. *The American Journal of Human Genetics* **76**, 663-671 (2005).
12. Birnbaum, R.Y. *et al.* Functional characterization of tissue-specific enhancers in the *DLX5/6* locus. *Human molecular genetics* **21**, 4930-4938 (2012).
13. Ramos-Zaldívar, H.M. *et al.* A novel description of a syndrome consisting of 7q21. 3 deletion including *DYNC1I1* with preserved *DLX5/6* without ectrodactyly: a case report. *Journal of medical case reports* **10**, 156 (2016).
14. Depew, M.J., Lufkin, T. & Rubenstein, J.L. Specification of jaw subdivisions by *Dlx* genes. *Science* **298**, 381-385 (2002).
15. Heude, É. *et al.* Jaw muscularization requires *Dlx* expression by cranial neural crest cells. *Proceedings of the National Academy of Sciences* **107**, 11441-11446 (2010).
16. Cole, J.B. *et al.* Genomewide association study of African children identifies association of *SCHIP1* and *PDE8A* with facial size and shape. *PLoS Genet* **12**, e1006174 (2016).
17. Shaffer, J.R. *et al.* Genome-wide association study reveals multiple loci influencing normal human facial morphology. *PLoS Genet* **12**, e1006149 (2016).
18. Lee, M.K. *et al.* Genome-wide association study of facial morphology reveals novel associations with *FREM1* and *PARK2*. *PLoS One* **12**(2017).
19. Staley, J.R. *et al.* PhenoScanner: a database of human genotype–phenotype associations. *Bioinformatics* **32**, 3207-3209 (2016).
20. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nature genetics* (2016).
21. Takimoto, A., Mohri, H., Kokubu, C., Hiraki, Y. & Shukunami, C.*PAX1* acts as a negative regulator of chondrocyte maturation. *Experimental cell research* **319**, 3128-3139 (2013).
22. Pohl, E. *et al.* A hypofunctional *PAX1* mutation causes autosomal recessively inherited otofaciocervical syndrome. *Human genetics* **132**, 1311-1320 (2013).